Interactive comment on wes-2023-30: A decision tree-based measure-correlate-predict approach for peak wind gust estimation from a global reanalysis dataset

S. Kartal, S. Basu and S. J Watson

May 10, 2023

General comments:
The manuscript titled "A decision tree-based measure-correlate-predict approach for peak wind gust estimation from a global reanalysis dataset" by S. Kartal et al. submitted for publication in Wind Energy Science adopts a new decision tree (DT)-based Measure-Correlate-Predict (MCP) approach called INTRIGUE that utilizes several meteorological variables from a public-domain reanalysis dataset to generate long-term, site-specific peak wind gust (W_p) estimation. The introduced four different models perform very well for nominal (non-extreme) conditions but the full potential for extreme conditions is yet to be seen. In particular, the application of this approach in the context of wind energy is very interesting, as wind gusts are critical for the highly fluctuating power production. The work presented here may serve as yet another approach to address this open issue.

I am happy to recommend the manuscript for publication in Wind Energy Science, as the subject and the approach used in the paper will be of broad international interest. I also strongly encourage publishing the code on github. My main remarks are listed below.

Specific comments:

- **Section 7.1 Self-Prediction:** Although the main ideas have been introduced, the overall presentation of this section still needs improvement. In particular, please clarify which data are used as training data and which are used for prediction. In particular, reconcile this section with Section 6.1 "Strategy for Splitting of Available Data".

- **Table 3-5, 7:** Highlight the essential items in the tables, such as the higher correlation of the means, or choose another more illustrative form of presentation.

- **Figure 6:** In addition to or instead of the confusion matrix, present the respective receiver operating characteristic curve.

- **Section 8 Limitations of the INTRIGUE Approach:** Add a discussion about which input features would most likely enable the INTRIGUE approach to predict wind gusts for extreme conditions.

Technical comments:

- **Line 3:** Define short & long-term in the manuscript.

- **Line 3:** Define nominal & extreme conditions in the manuscript.

- **Line 27:** The sentence: "The focus of the current... approach." might be better linked to the sentence: "In this paper, we propose a... reanlysis dataset." (Line 79).

- **Line 55 - 69:** This section could better follow directly after the section that ends at line 27.

- **Table 1:** Highlight most important input features for the approach

- **Figure 2, 7:** Add a liner relationship to each of the graphs to illustrate the correlation.

- **Figure 5:** Add Definition of W_{p10^m} in the caption

- **Figure 6:** Are the Confusion Matrix actually the exact same for the XGBoost and RF models?

- **Line 355:** Add the note that for the cross-predictions the different sites across the world should experience similar regional conditions.